An FPGA-Based Implementation of the CMS Global Calorimeter Trigger.

J. J. Brooke, D. G. Cussans, G. P. Heath, **D. M. Newbold**

H. H. Wills Physics Laboratory, University of Bristol,
Tyndall Avenue, Bristol BS8 1TL, UK
dave.newbold@cern.ch

A. J. Maddox, P. D. Rabbetts

Rutherford Appleton Laboratory, Chilton,
Oxfordshire OX11 0QX, UK.

Abstract

We present a new design for the CMS Level-1 Global Calorimeter Trigger, based upon FPGA and commodity serial link technologies.

For each LHC bunch-crossing, the GCT identifies the highest transverse energy electron, photon and jet candidates; calculates scalar and vector total transverse energies; performs jet-counting, and provides real-time luminosity estimates. The pipelined system logic is implemented using 0.18μm Xilinx FPGAs. The traditional system backplane is replaced by fast serial links for trigger data, and Ethernet for control. These technologies allow an improvement in system flexibility and a considerable reduction in cost, complexity and design time compared to an ASIC/VME-based solution.

I. THE CMS LEVEL-1 TRIGGER

The CMS Level-1 trigger [1] uses a subset of the data from the muon detectors and calorimeters in order to select bunch-crossings containing interesting events at a rate of no more than 100kHz. The Level-1 latency is required to be less than 128 LHC bunch-crossings (3.2μs) in order to limit the length of readout buffers. No significant deadtime is allowed; the system is therefore implemented in custom hardware, using fully pipelined logic synchronised to the LHC bunch clock.

The trigger algorithms are based upon the identification of trigger objects (μ, e/γ, jet, τ) and on measurement of missing and total transverse energy. Trigger objects are identified in the muon and calorimeter systems separately, and a subset of identified objects is selected from each according to transverse energy and other criteria. The full information describing the energy and position of the selected objects is sent to the global trigger, which combines information from the muon and calorimeter systems, and makes a Level-1 trigger decision based upon object energies and event topology.

The Level-1 trigger system is shown in outline in Figure 1. The calorimeter trigger consists of the following subsystems:

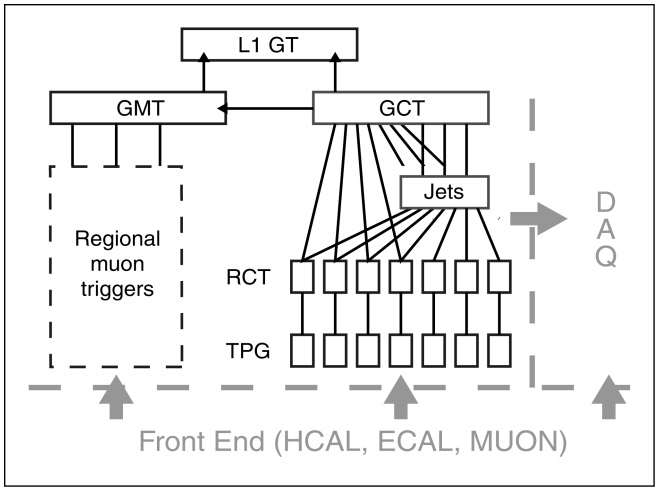

Figure 1: The CMS Level-1 trigger system

- The Trigger Primitive Generator (TPG) system, which processes the digitised calorimeter signals to produce coarser-grained trigger input data.
- The Regional Calorimeter Trigger (RCT) [2], which implements e/γ, jet and τ identification algorithms and calculates local transverse energy sums, in each of several separate detector regions.
- The Global Calorimeter Trigger (GCT), which fulfils the functions described below.

A first conceptual design for the GCT system has been already been presented [3]. Since the initial design study, the functional and technical requirements for the GCT and the Level-1 trigger system as a whole have been extended and refined, and new technologies have become available. We present here an updated design in the light of these changes.

II. GCT FUNCTIONALITY

The GCT is the final component in the calorimeter trigger chain. Its purpose is to implement the stages of the trigger algorithms that require information from the entire CMS calorimeter system. The GCT receives trigger object data from the RCT, performs several stages of data processing, and sends a reduced amount of data to the Global Trigger (GT). The functions of the GCT are explained in more detail below.

A. Trigger Object Sort

The principal function of the GCT is to reduce the number of trigger object candidates that need be considered by the GT. The is achieved by sorting the trigger objects identified by the RCT according to rank, and forwarding to the GT only a fixed number of objects. The rank of an object is in general based upon its transverse energy.

The RCT and GCT jet-clustering subsystem together identify objects of five different classes: e/γ, isolated e/γ, central jet, forward jet and τ. The GCT sort subsystem receives up to 72 objects of each class, and passes four of each class to the GT. Each trigger object is represented by a data word containing information about its rank and the location at which it was detected.

B. Jet Clustering

In the CMS Level-1 trigger, a jet candidate is effectively a localised hadronic deposit above threshold. Identification of such deposits is carried out by the RCT by summing the total energy within each cell of a fixed-size grid of calorimeter regions. However, improved trigger efficiency can be obtained by performing clustering of the identified deposits. The clustering algorithm currently under study consists of a further sum-and-compare operation on each possible 3x3 'window' of calorimeter regions [4].

The jet-clustering subsystem of the GCT receives a transverse energy sum for each calorimeter region, along with pattern bits indicating whether the deposit in each region is compatible with the decay of a τ to a single hadron. It performs the cluster algorithm over the whole calorimeter space, classifies the resulting jet candidates as central jet, forward jet or τ, performs an initial sort, and forwards the surviving objects to the final sort subsystem.

C. Jet Counting

In order to improve the trigger efficiency for rare multi-jet events, a jet multiplicity trigger is implemented alongside the main jet algorithms. The GCT determines the overall number of jet and τ objects fulfilling each of several sets of rank and position criteria; the totals are sent to the GT.

D. Global Energy Summation

The GCT is required to calculate the total transverse energy, and the magnitude and azimuthal direction of the missing transverse energy vector, for each bunch-crossing.

The calculation is based upon the transverse energy sum for each calorimeter region calculated by the RCT. The GCT weights each energy sum according to the angular position of the corresponding calorimeter region, and integrates them to give the two orthogonal components of the total transverse energy. Finally, the total and missing transverse energy are calculated from the components, and sent to the GT.

E. Luminosity Monitoring

Accurate knowledge of the LHC luminosity at the CMS interaction point is necessary in order to measure physics cross-sections. In addition to offline absolute luminosity measurements, it is also important to perform frequent and continuous measurement of relative luminosity. Since the GCT receives data from the whole calorimeter system for every bunch-crossing, it is possible to provide online luminosity monitoring on a bunch-by-bunch basis.

The choice of measurement technique is under study, but is likely to be based on rates of high transverse energy jets and/or global energy flows; this will make direct use of the jet counts and summary data from the global energy calculation subsystem. The calculated bunch luminosities are sent to the CMS detector control system at regular intervals.

F. Trigger Data Capture

The GCT provides information to the CMS DAQ system for every bunch-crossing that results in a Level-1 trigger. This information is used for online performance monitoring and fault detection, offline calculation of trigger efficiencies, and to identify regions of interest in the early stages of the Higher Level Trigger algorithms.

G. Control, Test and Monitoring

Automated control, test and monitoring of the GCT is an important requirement. Since there will be no physical access to the GCT system during LHC running, it must be possible to perform system setup, reset, test and reconfiguration without manual intervention. The GCT system is capable of performing rigorous self-test on a chip, board and system level whilst *in situ*, using test patterns and real or simulated physics data. Test and diagnostic data may be inserted or captured at many points in the system during normal running or setup.

III. GCT IMPLEMENTATION

A. System Overview

The main design goals for the GCT implementation are:

- Modularity, to simplify design, test and maintenance.
- Flexibility and room for expansion.
- Comprehensive self-test and monitoring capability.
- Low latency (<400ns)
- Reasonable cost.

The hardware design makes use of two key off-the-shelf technologies. Processing functions are performed using fast 0.18μm FPGAs. This approach results in a large reduction in system cost, since small production runs of a number of ASICs would otherwise be required. The use of programmable logic also allows excellent flexibility and modularity, since identical hardware may be used for a variety of processing functions. Secondly, all data transfer within and

out of the system is performed using high-speed serial links. The elimination of large numbers of parallel connections on cables and backplanes enables a reduction in complexity and an increase in system density.

The GCT system is located within the CMS underground electronics cavern, and is not subject to high radiation levels. The main GCT functions are implemented using only two types of board: data processing is performed by 14 identical Trigger Processor Modules (TPM); input data from the RCT are received, synchronised and reformatted by 14 Input Modules (IM). Each TPM and IM is implemented on a single 9U x 400mm board. System setup, control, test and monitoring are performed by an embedded CPU located on each module.

B. Input Module

Input data are transmitted from to the GCT as 80Mbit/s parallel differential ECL signals, carried on 108 20m copper cables. The total data flow into the GCT is around 250Gbit/s. The use of parallel signalling has several advantages; however, it requires a large amount of physical space for connectors and cabling. A dedicated input system is therefore used to synchronise and reformat the parallel data and retransmit it on short-haul high-speed serial links. This approach allows enough data to be received onto a board that each GCT subsystem can be accommodated on a single TPM. Input and algorithm processing functions are also effectively decoupled, allowing the use of the 'generic' TPM with fixed I/O capability for all algorithm functions.

Each IM receives data from eight input cables, terminated with 68-way SCSI2 connectors. The input synchronisation logic is shown in Figure 2. Programmable-length 320MHz and 80MHz FIFOs are used to align the incoming data to the GCT system clock and to the correct bunch-crossing. In order to ensure a robust link between RCT and GCT, no assumptions are made about the timing of the input data. The IMs may be programmed to correctly accept data of arbitrary phase on any input bit. Synchronisation setup is fully automated, and is carried out using test patterns at startup.

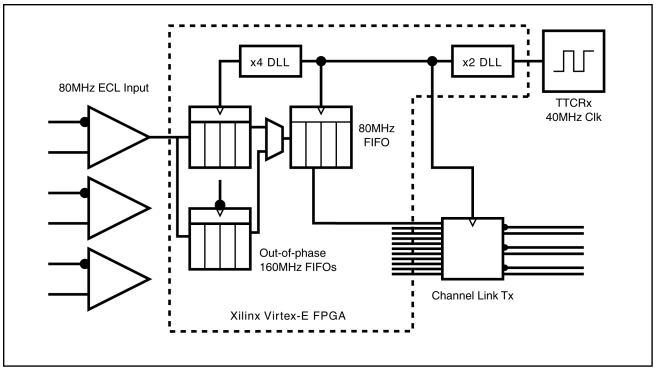

Figure 2: GCT input synchronisation logic.

Data are sent from the IMs to TPMs on short (<1m) 2.3Gbit/s serial links. These are implemented using a National Semiconductor Channel Link chipset [5], which serialises 28 bits of data at 80MHz. ITT/Cannon MDSM connectors are used [6], along with Amphenol Skewclear cable [7]. The mapping of data from ECL inputs to serial outputs is flexible. The output bandwidth of each IM is greater than the input bandwidth, to allow for the data duplication required by the jet cluster algorithm.

All synchronisation and processing functions are carried out within Xilinx Virtex-E FPGAs [8]. Each FPGA also contains a control interface, deep internal memory buffers for capture and playback of test data, and a derandomiser FIFO which interfaces to the board DAQ bus for capture of trigger data. The device under consideration here is the XCV300E.

All other circuitry (power, clock, control, etc) is common with the TPM; this allows a large reduction in design effort and testing time. Moreover, the IM and TPM boards are slot-interchangeable, to allow maximum flexibility in the way the system is laid out and to allow for future expansion.

C. Processor Module

The TPM provides a generic processing platform upon which a variety of trigger algorithms may be implemented. All TPMs share the same hardware design. A conceptual block diagram of the TPM is shown in Figure 3; certain details of the hardware implementation (e.g. number and placement of connectors) are still under study. Since the GCT is principally a data reduction system, the TPM is based upon a two-level logic 'tree', with large data input capacity and reduced data output capacity. Three 'Stage A' algorithm blocks receive data from the IMs or from other TPMs, and feed results to a single 'Stage B' algorithm block via wide data busses. Each block is implemented using a single Xilinx Virtex FPGA; the devices under consideration are the XCV1600E for Stage A and the XCV1000E for Stage B.

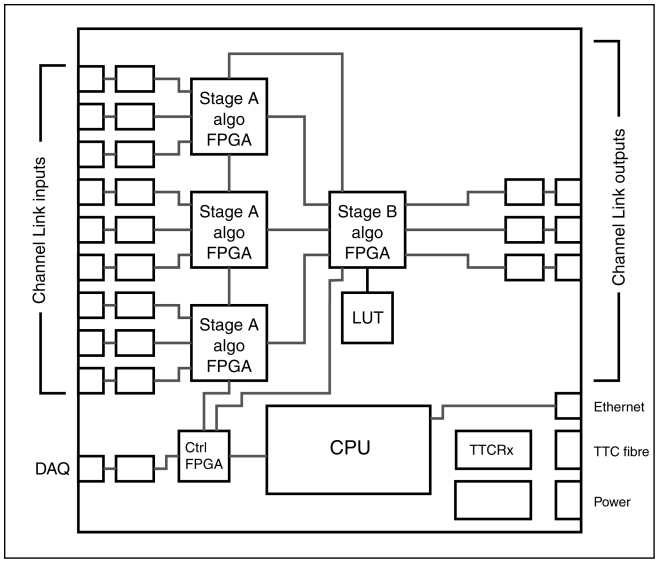

Figure 3: Trigger Processor Module block diagram.

In addition to the algorithm processing functions, each FPGA contains memory buffers and an interface to the board control and DAQ busses. This allows test data to be captured

and inserted into the processing chain, and trigger data to be captured and read out via the DAQ interface. The control and DAQ busses are implemented as a single unidirectional synchronous bus which travels point-to-point between the FPGAs, and is time-multiplexed between the two functions. Bus control is carried out by a separate control FPGA, which also implements the readout controller, TTCRx interface and a variety of other 'glue' functions.

All data transfer into, out of, and between TPMs is carried out over identical serial links. This includes the output links to the GT, though in this case the longer link distance may dictate that the chipset operates at 40MHz. All links are point-to-point cables; no system backplane is used for either data communication or control. Each TPM has 18 input links and 6 output links, plus two links for connection to the DAQ concentrator. No synchronisation is necessary at the TPM inputs, since any link delays are compensated for by adjusting the serialiser clock phase on the transmitting IM or TPM.

Clock and synchronisation signals are received by a TTCRx ASIC [9] on every TPM and IM. A passive optical fan-out within the GCT rack is used to distribute the incoming clock signals. All modules run with the same clock phase; the timing is adjusted during system setup using the fine-deskewing facility of the TTCRx.

Control and monitoring of the TPM is carried out by an embedded CPU. This is likely to be a commercial daughterboard; the device under consideration is the 486Core 'credit-card PC' from CompuLab [10]. This board contains a 486-compatible chipset, Ethernet interface and up to 32MB of flash memory. The use of an embedded CPU has several advantages over a backplane control bus such as VME. There is a reduction in cost, at the expense of bandwidth, and the local storage of FPGA program files is facilitated by the availability of a large non-volatile memory. The control CPU provides setup functions at system startup, and may be used to monitor the operation of the TPM during normal running by examining data captured at various points in the processing chain. High-level control and configuration is provided by the CMS trigger control software [1]. The CPU interfaces to the fast trigger logic via the control FPGA.

D. Algorithm Implementation

In this section, details are given of how the GCT algorithms map onto the resources provided by the TPMs. All algorithm logic is run at 80MHz clock speed, and is fully pipelined. Communication between FPGAs takes place at 80Mbit/s; this may be increased to 160Mbit/s after testing.

1) Trigger Object Sort

The e/γ and isolated e/γ sort algorithms involve the identification of the four highest rank objects from 72, and require one TPM each. Since the amount of logic required to perform a fast sort grows rapidly with the number of input items, the sort is split into three sequential stages. Groups of four objects are first presorted; no data is lost at this stage, but the presort reduces the complexity of logic in the following stages. The three second stage blocks find the four highest-rank objects from 24. The single third stage block finds the four overall highest-rank objects from the remaining 12. Each Stage A algorithm FPGA contains one second stage block, plus presort, and the Stage B FPGA contains the third stage block. If each object has six bits of rank data, as foreseen, the total latency of the sort is 14 cycles (seven bunch-crossings) [11].

The jet and τ sort trees are implemented in a similar way, but require only half a TPM (i.e. a single Stage A FPGA) apiece, since there are half as many input data, and the data are presorted by the jet cluster TPMs.

2) Jet Cluster

The cluster algorithm requires a large amount of logic, and occupies up to nine TPMs. The challenge in implementing the cluster algorithm is the requirement to share data between TPMs, since, however the calorimeter space is split up and allocated to modules, a large number of the 3x3 sum windows will lie across some boundary.

In order to simplify the system, there is no direct data sharing between TPMs. Instead, data duplication is performed by the IMs. If data from a calorimeter region are required by more than one TPM, then the data are duplicated and sent to both in parallel. This minimises the output bandwidth required on a TPM, and corresponds well to the resources required by other algorithms; however, increased IM output bandwidth is needed. There are advantages to processing a 'strip' of calorimeter regions in each TPM. This actually maximises the amount of data duplication, but allows a reduction in the complexity of the algorithm logic. However, there are several other workable solutions [12].

Each jet cluster TPM sends output data to the central jet, forward jet and τ sort modules, and to the global energy sum module.

3) Global Energy Sum

The calculation of total and missing transverse energy requires two Stage-A FPGAs, and so can coexist on the same TPM as the τ sort. The energy sum logic receives the two orthogonal components of transverse energy, integrated over the area covered by each jet cluster module. The components are summed, and used to calculate the final total and missing transverse energy sums. The final stage of the calculation requires the use of a large external lookup table. The number of energy bits used at the various stages of the energy summation is chosen such that rounding errors do not contribute significantly to the precision of the final sums.

4) DAQ / Luminosity

A single TPM acts as DAQ concentrator and derandomiser. It receives data from other TPMs and IMs upon a Level-1 Accept; performs derandomisation; packages the data into a standard format; and interfaces to an SLINK driver for communication with the CMS DAQ system. The luminosity monitor is parasitic on the DAQ logic; by examining the DAQ data stream, it can access any part of the

captured trigger data. This requires that some trigger data be sent for every bunch-crossing, regardless of the presence of a Level-1 Accept. This is acceptable, as there is no requirement for the data to be sent synchronously, it can be multiplexed with the normal DAQ data flow into the TPM. The luminosity logic consists largely of counters to integrate the rate of high Et, low-h jets and other selected physics objects.

IV. TESTS AND PROTOTYPES

A. Processing Tests

The GCT design depends upon the ability to carry out a variety of algorithmic functions in fully pipelined logic using FPGAs. In order to demonstrate the feasibility of this approach, a programme of prototyping has been carried out, which will culminate in the construction of a complete prototype TPM vertical slice in 2001.

In order to gain experience with the Xilinx Virtex family, a first FPGA test platform was constructed in 1999, and used to test the sort algorithm. An XCV300-5 FPGA was fed with pseudorandom data from hardwired LFSRs. Due to the limited logic capacity of the XCV300, only the second-stage sort was implemented, finding the four highest-rank objects from 24. The algorithm performed correctly up to a clock speed of 85MHz, compared to a design speed of 80MHz, thus demonstrating the feasibility of the sort implementation with the technology current at that time. This result was in full agreement with a software simulation.

A second technology evaluation platform, based around XCV1000E and XCV1600E FPGAs, along with external SRAM, is now under construction. This will allow full *in silico* testing of all final GCT algorithms prior to the construction of a prototype TPM. The board will also be used to test the TTC and embedded CPU devices.

B. Data Link Tests

The correct operation of the calorimeter trigger depends upon the ability to correctly synchronise systems separated by some distance, and to move large amounts of data between them with very low error rates. The GCT uses two link technologies: parallel 80Mbit/s ECL and high-speed serial LVDS. A series of tests on these technologies have been carried out in order to prove the feasibility of the concept, to measure the achievable bit-error rates under various conditions, and to optimise the choice of chipsets, cables and connectors, and synchronisation scheme.

A link test board was built in 1999 to enable a first set of tests of bit-error rates. The board contains a Xilinx XC4020XLA FPGA which acts as data source and sink, a control interface to a PC, and full-speed ECL and Channel Link input and output interfaces. Acceptable combinations of cables and connectors were identified in both cases [1]. Bit-error rates under lab conditions, and using pseudo-random data, were found to be less than 10^{-12} for the ECL interface and 10^{-14} for the Channel Link, both well within acceptable limits for correct operation of the trigger system. Eye diagrams observed during the tests are shown in Figure 4.

A second link test board is being designed, which will implement the full functionality of an IM for a single channel. This will allow testing of the automatic synchronisation technique, and together with the new FPGA test board will constitute a full vertical slice of the GCT system.

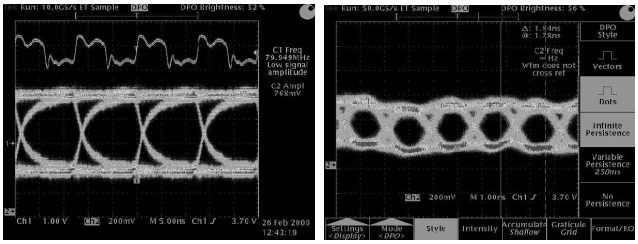

Figure 4: Eye diagrams obtained during GCT link technology tests. Left: 80Mbit/s ECL. Right: 560Mbit/s LVDS.

V. CONCLUSION

The CMS GCT design has been updated in the light of new requirements and improvements in available technology. The feasibility of the new approach has been demonstrated in hardware, and the project is now ready to move to the prototyping phase. The use of off-the-shelf technologies and high-speed programmable logic has resulted in an improvement in system flexibility and robustness, whilst allowing a reduction in design time and cost.

VI. REFERENCES

[1] CMS Level-1 Trigger Technical Design Report, in preparation.

[2] Contribution of W. Smith, these proceedings.

[3] "The Global Calorimeter Trigger for CMS", D. Bailey *et al*, 4th Workshop on Electronics for LHC Experiments, Rome, CERN-LHCC-98-036.

[4] CMS Note 2000/055, S. Eno *et al*.

[5] http://www.national.com/ds/DS/DS90CR287.pdf

[6] http://media.ittcannon.com/pdf/catalogs/MDSM.pdf

[7] http://www.spectra-strip.amphenol.com/

[8] http://www.xilinx.com/products/virtex.htm

[9] http://ttc.web.cern.ch/TTC/intro.html

[10] http://www.compulab.co.il/486core.htm

[11] CMS Note 2000/Draft, "Sort processing in the Global Calorimeter Trigger", J. Brooke *et al*.

[12] CMS Note 2000/draft, "Jet Clustering in the Global Calorimeter Trigger", J. Brooke *et al*.